Appendix C

Reference for Paired T tests

Using the *t*-test for paired measurements

For a comparison of more than one pair of split sample test results, ***the t*-test for paired measurements** can be used. Typically, this test uses the difference between each of the paired tests and determines whether the result is statistically different from 0, which is the expected theoretical result when testing separate portions of a single product. Thus it is the difference *within* pairs, not between pairs, that is being tested.

The form of the *t*-test for paired measurements is:

$$t = \frac{\left\| \overline{X}_d \right| - \left| X_O \right\|}{\frac{s_d}{\sqrt{n}}} \qquad \text{Eq. (1)}$$

Where: $\overline{X_d}$ = Average of the differences between the individual split sample test results

X_O = The value of the expected difference between split sample tests. In most cases $X_0=0$, as there is no expected difference between contractor and agency tests. However, X_0 could reflect an expected correlation value between testing entities.

s_d = Standard deviation of the differences between the split sample test results

n = Number of split samples

The calculated t value is then compared to the critical value, t_{crit}, obtained from the appropriate table with n-1 degrees of freedom. Use Table 1 for a one tailed analysis and Table 2 for a two tailed analysis. The majority of paired-*t* evaluations will utilize a two tailed distribution as the acceptance criteria reflect both upper and lower limits.

One of two outcomes will be considered:

1. $t < t_{crit}$ - Therefore the tests are accepted, as the differences between the test results are statistically likely to occur.

2. $t \geq t_{crit}$ - Therefore the tests are not accepted, as the differences between the test results are greater than is statistically likely to occur.

The data should be analyzed with the same number of decimal places as used for the acceptance testing. For example, if gradation numbers are normally reported in whole numbers, then use whole numbers for the paired t test. If the 200 sieve is normally reported to the 10th decimal place, for example 6.5% then use 6.5 for the paired t test.

Example 1

Verification Using Paired *t*-test

(one tailed distribution: i.e., upper or lower limit, α = 0.10)

Consider the following results of split sample testing for concrete compressive strength. Use a paired t-analysis to determine whether a difference exists between the contractor and agency results.

Sample	Contractor	Agency	Difference
1	5025	5360	-335
2	5270	5530	-260
3	4865	4915	-50
4	4950	4890	60
5	5120	4430	690
6	4750	4805	-55
7	5145	4985	160
8	4890	5525	-635

$\overline{X_d}$ = -53.13

s_d = 391.21

Using equation 1:

$$t = \frac{\left\| -53.13 \right| - \left| 0 \right\|}{\frac{391.21}{\sqrt{8}}} = 0.384$$

Using a level of significance, α = 0.10

Degrees of Freedom = n-1 = 8-1 = 7

From Table 1, t_{crit} = 1.415

Conclusion: Since 0.384 < 1.415, we assume that the paired tests are the same. We therefore assume that the contractor's and agency's test results from paired measurements indicate that the test methods, technicians, and test equipment ***are*** providing similar results. Again, keep in mind that we can conclude nothing about the material.

Table 1 - Values for $t_{crictical}$
One Tailed Significance

One Tailed Significance						
Degrees of freedom (n-1)	**α = 0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**
1	3.078	6.314	12.706	31.821	63.657	318.300
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.214
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.305	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.611
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.704	3.307
60	1.296	1.671	2.000	2.390	2.660	3.232
80	1.292	1.664	1.990	2.374	2.639	3.195
100	1.290	1.660	1.984	2.365	2.626	3.174
4	1.282	1.645	1.960	2.326	2.576	3.090

Example 2

Verification Using Paired *t*-test

(two tailed distribution: i.e., upper and lower limit, α = 0.05)

Consider the following results of split sample testing for asphalt binder content. Use a paired t-test to determine whether a difference exists between the contractor and agency results.

Sample	Contractor	Agency	Difference
1	5.6	4.8	0.8
2	5.4	4.4	1.0
3	4.7	4.1	0.6
4	5.5	5.1	0.4
5	5.8	4.9	0.9
$\overline{X_d}$ =			0.74
s_d =			0.24

Using equation 1:

$$t = \frac{\left\| 0.74 \right| - \left| 0 \right\|}{\frac{0.24}{\sqrt{5}}} = 6.894$$

Using a level of significance, α = 0.05

Degrees of Freedom = n-1 = 5-1 = 4

From Table 2, t_{crit} = 2.776

Conclusion: Since 6.894 > 2.776, we assume that the paired tests are different. We therefore assume that the contractor's and agency's test results from paired measurements indicate that the test methods, technicians, and test equipment ***are not*** providing similar results. Keep in mind that we can conclude nothing about the material.

Table 2 - Values for $t_{crictical}$
Two Tailed Significance

Two Tailed Significance						
Degrees of freedom (n-1)	**α = 0.20**	**0.10**	**0.05**	**0.02**	**0.01**	**0.002**
1	3.078	6.314	12.706	31.821	63.657	318.300
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.214
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.305	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.611
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.704	3.307
60	1.296	1.671	2.000	2.390	2.660	3.232
80	1.292	1.664	1.990	2.374	2.639	3.195
100	1.290	1.660	1.984	2.365	2.626	3.174
4	1.282	1.645	1.960	2.326	2.576	3.090